Report No. 72-0037 Contract No. NAS8-27359 NTIS HC \$5.00

MULTI-PROCESSING SYSTEM STUDY

FINAL REPORT

(NASA-CR-124026) MULTI-PROCESSING SYSTEM STUDY Final Report (M&S Computing, Inc.) 60 p HC \$5.00 CSCL 22B

N73-16873

Unclas G3/31 16946

Prepared for:

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION George C. Marshall Space Flight Center Marshall Space Flight Center, Alabama 35812

M&S) OMPUTING, INC.

PREFACE

This document summarizes the results of a multi-processor systems design review. The results of this review are anticipated to be used for the design of a SUMC - Multi-Processing System.

This effort was performed for NASA-MSFC under Contract No. NAS8-27359, Modification No. S/A3.

Participants were:

- T. T. Schansman
- J. R. Conway
- E. I. Eastin

TABLE OF CONTENTS

Section			
1.	INTRODUCTION		
2.	IBM DESIGN EVALUATION	2	
	2.1 IBM Design - Version 1 - Review Summary	3	
3.	RCA DESIGN EVALUATION		
	3.1 RCA Design - Review Summary	26	
חחים מ	D ENIC ES	58	
REFERENCES			

1. INTRODUCTION

This report summarizes the results of a multi-processor systems design review. The systems were reviewed against use in a "Space Station Environment".

The purpose of the review was to evaluate the proposed designs from a systems viewpoint in general and from the systems software in particular. The recommendations resulting from this evaluation are anticipated to be considered for the design of a multiprocessing system built around a SUMC.

Two multi-processing system designs were reviewed; one design proposed by the IBM Corporation, one design as proposed by the RCA Corporation. It was the original intent to also review a design proposed by the Intermetrics Corporation. However, the Intermetrics design was not available in time for a full evaluation, and was therefore excluded from this report.

In general, the designs reviewed were highly functional and many questions could not be answered. However, several major issues were uncovered which could be evaluated to some detail, and which could greatly impact the SUMC-MP design.

The major issues relevant to a multi-processing system design revolve around the following functions:

- 1) Storage Management
- 2) Processor Management
- 3) Intermodule Communication
- 4) Memory Access Interference
- 5) System Efficiency
- 6) System Recovery/Reliability

Section 2 of this document summarizes the results of the IBM design review. Section 3 summarizes the results of the RCA design review. References are provided at the end of this report.

2. IBM DESIGN EVALUATION

Two designs were, in effect, submitted by IBM. Paragraph 2.1 summarizes the evaluation of the first design, Paragraph 2.2 summarizes the evaluation of a second design, an amended version of the first.

The first design contained many shortcomings and was considered the lesser of the two designs (IBM and RCA). The amended version improved the design considerably; however, it was far more functional than the RCA design and therefore difficult to compare.

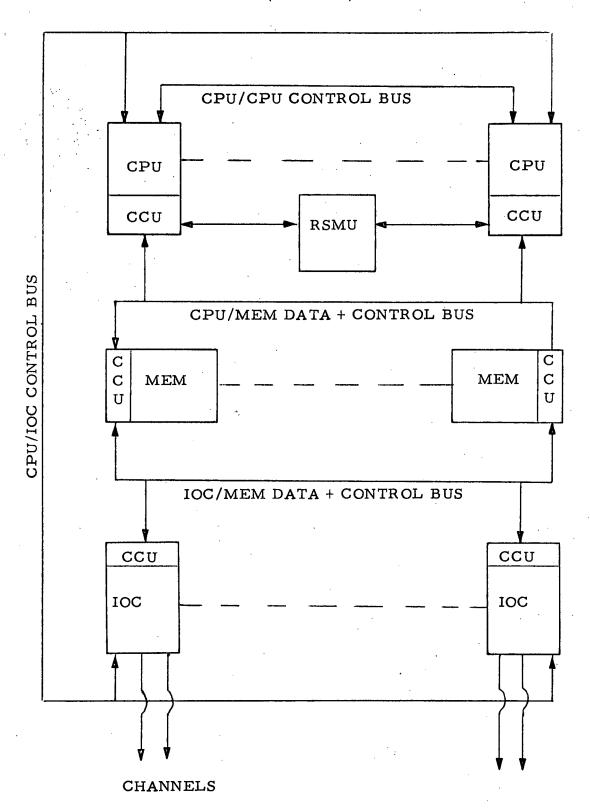
2.1 IBM DESIGN - VERSION 1 - REVIEW SUMMARY MAIN TOPICS

- SYSTEM OVERVIEW
- ✓ CONFIGURATION CONTROL UNIT
 - o Address translation
 - o Communication control
 - o State control
- ✓ RECONFIGURATION STATE MONITOR UNIT
- INPUT/OUTPUT
- ✓ MULTIPROCESSING INSTRUCTIONS
- ✓ FAULT TOLERANCE
- ✓ SUMMARY

✓ SYSTEM OVERVIEW

- Processor architecture is S360 based but not compatible.
- o High emphasis on automatic partitioning into independent systems without physical disconnects.
- o High emphasis on not affecting SUMC design (as it existed at that time).
- Floating Executive intended.

(continued)



CONFIGURATION CONTROL UNIT (CCU)

Purpose:

- o To allow logical addressing of memory modules ATU.
- o To allow partitioning of modules into independent systems, without physical disconnects. Communication control.
- o To allow controlled loading, initializing, reconfiguration, etc., State control.

Implementation Notes:

- o Each module has a CCU.
- o Special instructions provided for CCU control.

ADDRESS TRANSLATION UNIT

Purpose:

- o To allow physical modules to be replaced without affecting software addresses.
- o To assign preferential storage addresses (0-n) to physical locations.

Functional Characteristics:

- o Each CPU has an ATU.
- o Each CPU may address any module.
- o ATU may be used to prevent access to physical modules (partitioning into independent systems).

ADDRESS TRANSLATION UNIT (continued)

Comments:

- o Re-assignment of memory module addresses can be more cost effectively performed in the memory modules (HUGHES-ARMMS).
- o Design will be dictated by selected paging mechanism.
- by paging CPU's address same logical addresses.

PREFERENTIAL STORAGE AREA

Purpose:

- o Contains information necessary to perform CPU initialization.
- o Contains save areas for interrupt handling.
- o Contains I/O communication control information.
- o Used for diagnostic logouts.

Basic Need:

- o Each CPU needs to generate identical fixed logical addresses for various communication purposes.
- o Logical addresses need to be transformed to physical addresses unique to each CPU.

PREFERENTIAL STORAGE AREA (continued)

S360

0	Initializa	ation			
`		-	Initial program load PSW		
		-	Initial program load CCWl		
		-	Initial program load CCW2		
0	Interrup	t/Trap	Control		
		· .	External old PSW		
			Supervisor call old PSW		
		-	Program check old PSW		
			Machine check old PSW		
		-	I/O old PSW		
		-	External new PSW		
		· •	Supervisor call new PSW		
		-	Program check new PSW		
		-	Machine check new PSW		
		-	I/O new PSW		
0	I/O Con	trol			
		-	Channel status word		
			Channel address word		
o	Timer				
			Timer		
o	Diagnos	tics			
		_	Model dependent		

COMMUNICATION CELL ADDRESSING SUMC/MP

- o Special Translation Unit undesirable
 - Translation unit used for paging.
 - Translation unit may not be unique for each CPU.
- Other approaches
 - Generate unique logical (i.e., paged) address internal to CPU, based on (externally) assigned identity.
 - Use pointer-addressed common communication
 area, put CPU identity in communication cell.
 Locking of (shared) pointer(s) may cause problems.

• COMMUNICATION CONTROL

Purpose:

- To logically disconnect modules from each other to form independent systems for:
 - Maintenance
 - Checkout
 - Debug (?)
 - Bring up new experiments (?)

Functional Characteristics:

o Each module has a communication state register to inhibit unwanted communications.

	•	CPU.	IOC	MEM
	MEM	15	15	-
to	CPU	4	-	
from	CPU	4	4	4
to	IOC	3	· -	-
from	IOC	3	_	3

o Registers are controlled through special instructions.

COMMUNICATION CONTROL

(continued)

Comments:

- o Need for debug or independent systems questionable system always needs protection from software errors.
- o Seems very useful for maintenance and checkout, which may bypass other system protection features.

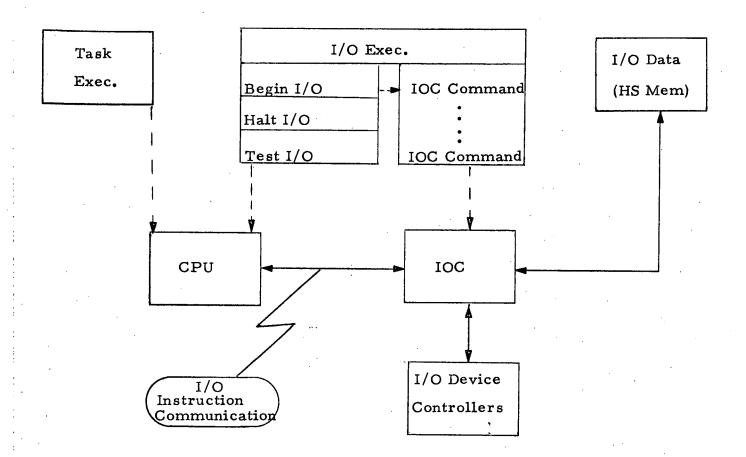
STATE CONTROL

Purpose:

o To enable controlled initialization, reconfiguration

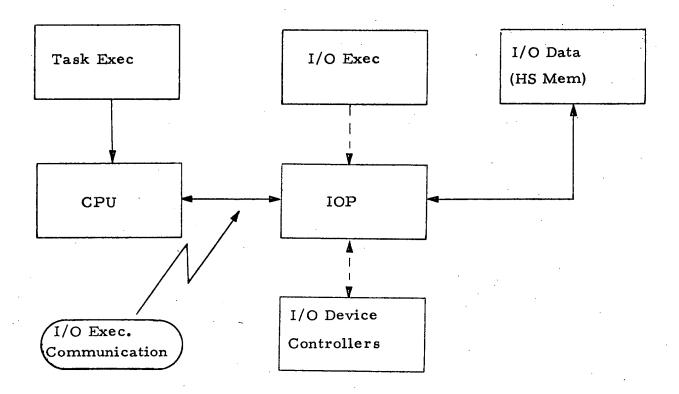
Comments:

o Necessary, but is driven by other design features



- o Conventional (360) approach
- o IOC is pure channel control unit
- o I/O termination interrupts handled by specific (initiating) CPU
- o Vehicle bus controller likely to be separate, and more intelligent, unit.

INPUT/OUTPUT (continued) RCA DESIGN



- o Execution of I/O Exec and Task Exec physically separated.
- o IOP is (near) full-fledged processor.
- o I/O termination signals can be decoupled from specific (initiating)

 CPU.
- o I/O Exec integral part of IOP design.
- o Vehicle bus controller may be regular IOP.

✓ RECONFIGURATION STATE MONITOR UNIT

(RSMU)

Purpose:

To detect major malfunctions during reconfiguration

Functional Characteristics:

- o Detects multiple state 3's (reconfiguration state).
- o Detects excessive time in state 3.
- Activates alarms and generates an interrupt to the CPU's.

Implementation Characteristics:

Duplexed Hardware

Operational Characteristics:

o Initiated by CPU

RECONFIGURATION STATE MONITOR UNIT

(RSMU)

(continued)

General Comments:

- o Special emphasis on reconfiguration state not justified.
- o Performs standard "watchdog" service necessary for any running task (exception detection of multiple state 3's).
- o Reconfiguration is a manually initiated and controlled operation.

Recommendation:

- o Drop further consideration of this unit.
- o Use standard external interval timer to provide watchdog service.

✓ MULTIPROCESSING INSTRUCTIONS

Load/Store CCU/CPU CCU/E

To allow control of configuration control registers and states of modules.

- o Test and Set
 - To perform locking of shared instructions or data.
- o Delay (n microseconds)

 Main use in conjunction with Test and Set.
- Store Identity

To be used with exec. interface.

FAULT TOLERANCE

- by redundant elements to aid in detection and isolation of faults".
- o Configuration Control Registers.
- o Reconfiguration State Monitor Unit.
- o Memory parity.
- o Storage protection lock and key.
- o Memory module addresses re-assignable (through ATU).
- o Memory access timeouts.

√ SUMMARY

- o Design is largely traditional
- DELAY instruction seems useful
- o PSA design is not fully worked out in RCA design
- o RCA-IOP design preferable probably

2.2 IBM - VERSION 2 - REVIEW SUMMARY NOTABLE DIFFERENCES FROM VERSION 1

INTERMODULE COMMUNICATION

Changed from busses (4) to communication matrices (2)

Comment: Seems better. Not enough info in handout to evaluate.

o PROCESSOR ARCHITECTURE

Changed to "augmented" S360

No comment

o PREFERENTIAL STORAGE ADDRESSING

Changed from address translate to control (base) register.

Comment: Good chance.

o VIRTUAL STORAGE

Added to system architecture

Comment: Use of "segments" and "reference bits" unclear.

IBM - VERSION 2

NOTABLE DIFFERENCES FROM VERSION 1 (continued)

o VIRTUAL MACHINE

Added to Executive Design

Comments: Sensible addition, as long as application programs can run on real machine in parallel with virtual machine. Approach not clear from handout.

RECONFIGURATION STATE MONITOR UNIT

Removed

Comment: Good

CONFIGURATION CONTROL UNIT

Automatic partitioning replaced by manual partitioning.

Comment: Good

MEMORY MODULE FEATURES

- Content correctable (?)
- Interleave
- Memory registers

Comment: Logical changes. Not sure above effectiveness of memory interleave in a paged system, but will not hurt.

IBM - VERSION 2

SUMMARY

- o Much improved over Version 1
- o No new features evident that should be considered in RCA design.
- o Virtual machine concept worthwhile addition to RCA Executive.

3. RCA DESIGN EVALUATION

The RCA design was documented in far more detail than the IBM design and thus subject to closer scrutiny. The design was certainly the more interesting of the two.

It is anticipated that the SUMC Multiprocessor System will be functionally closer to the proposed RCA design.

3.1 RCA DESIGN - REVIEW SUMMARY

GENERAL COMMENT

Emphasis of this evaluation is the <u>system</u> not the <u>software</u> complexity.

MAIN TOPICS

- ✓ Virtual Memory
- √ Hardware Task Switching
- ✓ Intermodule Communications
- ✓ Information Protection
- √ Fault Tolerance
- √ Potential SUMC Design Impacts

MAIN CRITICISMS

- ✓ CPU assignment to tasks too inflexible
- √ Paging impact on performance too large
- Emphasis on system throughput not compatible with real time aspects.

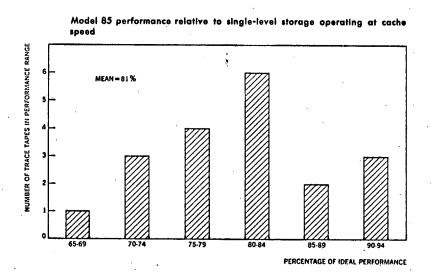
o ADVANTAGES

- Relieves application programmer from segmenting and overlaying programs.
- Allows application program execution to be started, even though insufficient high speed storage is available to contain whole program.

o DISADVANTAGES

- Potentially large impact on performance!!
- Slightly increases Control Executive/Hardware complexity.
- Requires "look-ahead" on Variable Field length and I/O instructions.

SAMPLE PAGING PERFORMANCE



Parameters

- o Speed ratio: Aux Mem/HS Mem ≈ 10
- o "Page" size: 64 bytes
- o Aux Mem path width: 16 bytes (1/4 page)
- o 4 way interleave
- o Effective transfer speed per Aux Mem word ≈ HS Mem word access time

APPLICABILITY TO SUMC/MP

- Application program > high speed store
 Not a common occurrence.
- Total application program storage > high speed store.

More common due to multiprocessing/ multiprogramming.

Conclusion: Main advantage on SUMC/MP is potentially more efficient high speed store utilization.

Therefore, better throughput.

Note: Better throughput can only be realized if page faulting and paging is kept to absolute minimum necessary.

PAGING

o GENERAL COMMENTS

- Currently most popular approach to Virtual

 Memory implementation.
- Subject to "internal fragmentation", thus offsetting potential storage efficiency somewhat.

o DESIGN PARAMETERS

- Fetch Policy: When is a page to be moved to high speed memory?
- Replacement Policy: Which page in high speed memory will be removed to auxiliary memory?
- Page Size

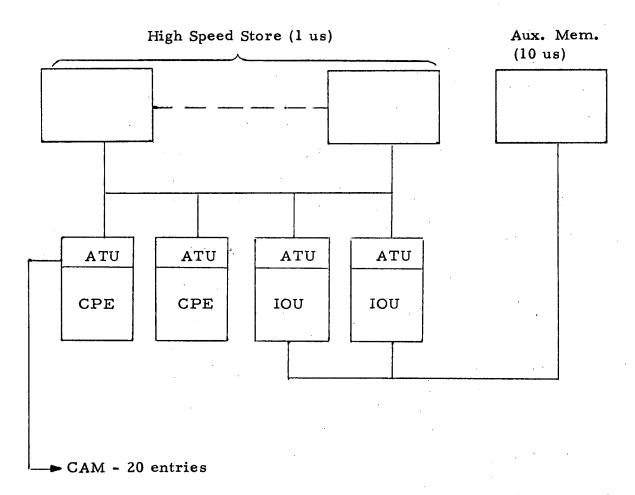
o MAIN IMPLEMENTATION PROBLEM

- Page address translate mechanism

RCA DESIGN

o	FETCH POLICY	-	Demand paging, except for
	•		executive.
0	REPLACEMENT	-	?
	POLICY		Locked pages.
o	PAGE SIZE	-	128/256 words. Due to selection
			of aux. storage parameters.
o	OTHER	-	Name space > virtual memory.
		-	Each task has translate table.

RCA DESIGN (continued)



10 reserved for Exec.

1280/2560 words directly addressable application storage

RCA DESIGN (continued)

o ADVANTAGES

- ATU search time only holds up one CPE.
- Application related info. protection may be included in ATU on a page basis.

o DISADVANTAGES

- ATU reload, for both CPE and IOU, is a common occurrence.

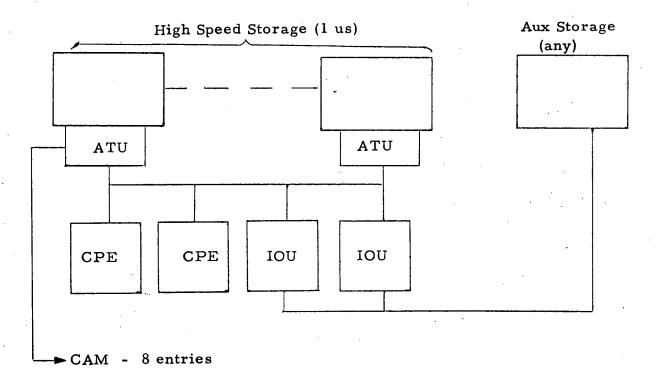
Requires additional memory accesses even though page is in high speed storage.

- Page swap potentially involves multiple CPE's/IOU's.
- Page size too small for disk/drum.
- Duplication of entries for Exec. pages.

VIRTUAL MEMORY

PROPOSED APPROACH

- o FETCH POLICY Demand paging with system-directed prepaging.
- o REPLACEMENT Least Recently Replaced (LRR)/Task or floating LRU
 POLICY Priority/Free Pages.
- o PAGE SIZE Min. 1000 words.



All high speed storage directly accessible

VIRTUAL MEMORY

PROPOSED APPROACH

o ADVANTAGES

- All high speed storage directly accessible by all CPE's/IOU's.
- Potential mix of paged and non-paged memories possible.
- ATU's are maintained without CPE participation.

o DISADVANTAGES

- ATU search holds up all CPE's/IOU's. However, search time short and probably overlapped with current memory cycle.
- Page fault indicated by time out.

/ HARDWARE TASK SWITCHING

RCA DESIGN

•			
	ROUTINE FUNCTION	PRIORITY LEVEL	
Systems Engineering	Power Failure Module Failure Entry Mode Switching MMU Code Check Failure	15 14 13 12	
Executive	Executive Entry Task Control I/O Request Page Fetch Interval Timer	11 10 8 8 7 8 7 11	
U ser	User Level 6 User Level 5 User Level 4 User Level 3 User Level 2 User Level 1 User Level 0	6 5 4 3 2 1 0 0 i.i.d.	

- o Each level has own set of registers.
- o Registers are never automatically stored in memory.
- o Tasks are pre-assigned to user levels.

/ HARDWARE TASK SWITCHING

RCA DESIGN (continued)

o GENERAL COMMENTS

- Performance advantage negligible with one exception.
- Pre-assigning tasks to CPE's is unsound policy.
- Not re-assigning waiting tasks to other CPU's is unsound policy.
- Most common need for store/restore subroutine usage not taken care of.
- However, retain large SPM and interrupt register

 Implement such that it can be used in different ways.

HARDWARE TASK SWITCHING PERFORMANCE EVALUATION

- I TO INITIATE A TASK UPON OCCURRENCE OF SPECIAL EVENT,

 BY JUST ENABLING LEVEL; I. E., REGISTERS DO NOT HAVE TO

 BE LOADED.
 - o System Engineering Levels
 - Occurrence highly exceptional.
 - Need only a few registers to start and probably execute.
 - Advantage negligible.
 - o Executive Levels
 - Occurrence regular.
 - Estimate average of 4 registers contain start information.
 - Some advantage.
 - o User Levels
 - Usage not applicable. The registers, if any,
 have to be loaded from memory upon assignment to CPU.
 - o No advantage.

PERFORMANCE EVALUATION (continued)

- II TO SAVE TASK INFORMATION UPON PRIORITY INTERRUPT,
 SUCH THAT IT CAN BE AUTOMATICALLY RESTARTED WHEN
 HIGHER LEVEL TASK IS COMPLETED.
 - o System Engineering Levels
 - Occurrence of one event exceptional, occurrence of multiple events near zero.
 - No advantage.
 - o Executive Levels
 - Levels cannot interrupt each other, with exception of intermodule communication.
 - Interrupt by Systems Engineering Level exceptional.
 - Advantage, at best, doubtful.
 - o User Levels
 - Suspension (necessary save of registers) is common occurrence.
 - Re-activation on same processor not a common occurrence.

HARDWARE TASK SWITCHING

PERFORMANCE EVALUATION

(continued)

- 1) Interrupt by SE Levels
 - Re-activation on same processor unlikely.
 - Occurrence exceptional.
 - No advantage.
- 2) Interrupt by Executive Levels
 - Re-activation on same processor normal for calls not causing "wait" condition.
 - Regular occurrence.
 - Advantageous.
- 3) Interrupt by User Level
 - Re-activation on same processor doubtful.
 - Could be regular occurrence.
 - No advantage.

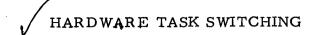
/ HARDWARE TASK SWITCHING

PERFORMANCE EVALUATION

(continued)

SUMMARY

- o Only advantage on certain Executive calls.
- o Number of levels and separate level registers can be decreased without affecting performance.
- o Re-assigning "suspended" tasks to other CPE incurs large overhead.
- o Subroutine call store/restore is not handled.



PROPOSED IMPROVEMENTS

(CPE)

o Priority Levels and Assignments

Mandatory Register Sets	Routine Function	Priority Levels
1 ·	System Engineering	TBD
1	Cross Module Comm.	1
1	Interval Timer	1
1	Exec. Call	1
TBD	Processing Trap	TBD
4-5	User Level	None

- O User level has 4-5 hierarchical levels invoked on subroutine call (local).
- o Additional instructions
 - CALL SUBROUTINE (local) .
 - STORE/RESTORE REGISTERS LEVEL X

INTERMODULE COMMUNICATION RCA DESIGN

o FUNCTIONAL CHARACTERISTICS

- CPU to CPU: None

- CPU to IOP: Stack - Global

IOP to CPU: Stack - Local

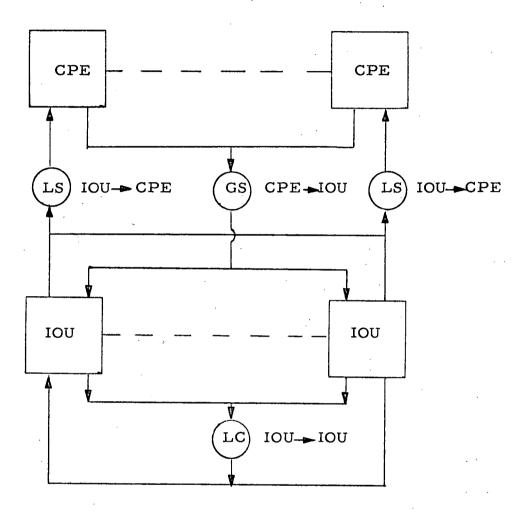
IOP to IOP: Single - Local

o IMPLEMENTATION CHARACTERISTICS

- All communication through fixed memory locations.
- Attention signal derived from continuous module polling of specific memory locations.
- Polling rate indeterminable.
- Polling performed by software.

INTERMODULE COMMUNICATION

RCA DESIGN



GS - Global Stack

LS - Local Stack

LC - Local Cell - common

✓ INTERMODULE COMMUNICATION REQUIREMENTS ANALYSIS

INPUT/OUTPUT SUPPORT - Solicited

- CPU \rightarrow (any) IOU

Purpose: Request for data transfer

Type: Global Request - Any IOU

Info: Command info.

Data address

Request ID (not CPU ID)

Request Priority

- IOU→(undetermined) CPU

Purpose: Signal Request completed

Type: Global - CPU assignment not known

to I/O manager

Info: Request ID

Request Status

INTERMODULE COMMUNICATION REQUIREMENTS ANALYSIS

o INPUT/OUTPUT SUPPORT - Solicited

- CPU→ (any) IOU - special case

Purpose: Feed/request data to/from continuous

I/O task (e.g., telemetry)

Type: Global - IOU assignment unknown

Info: Command Info

Data address

Evaluation Notes

- 1) Only CPU ID is known
- 2) Request status goes to specific CPE consequence of 1. Requires task to remain with specific CPE.

INTERMODULE COMMUNICATION

REQUIREMENTS ANALYSIS

o INPUT/OUTPUT SUPPORT - unsolicited

- $IOU \rightarrow (any) CPU$

Purpose: To signal that (uninitiated) external event

has occurred.

Type: Global - CPU assignment unknown.

Info: Command Info.

Data address, if any.

Evaluation Notes

- 1) Design requires signal to specific CPU.
- 2) This and previous case may be solved by IOU executing CPE Task Control.

✓ INTERMODULE COMMUNICATIONS REQUIREMENTS ANALYSIS

o CPU TASK ASSIGNMENT

- CPU→CPU

Purpose: To assign a (higher priority) task to an

active CPE.

Type:

Local - Specific CPE

Info:

Task control block

Evaluation Notes:

- 1) Not available in design
- 2) May be solved by IOU executing CPE task control.
- 3) May be solved by assigning (hardware)

 Executive Control to "lowest priority" CPU.

✓ INTERMODULE COMMUNICATION

REQUIREMENTS ANALYSIS

- o IOU TASK CONTROL
 - IOU→IOU

Similar to CPU task control

- Evaluation Notes

Single cell available for IOU→IOU communication

/ INTERMODULE COMMUNICATION

RCA DESIGN

COMMENTS

- o IOU REQUIRED TO PERFORM
 - Part of CPE task control
 - Intertask synchronization
 - Reconfiguration
 - IOU associated functions

Consequences

- No separation of I/O and CPE management
- Potential bottleneck if one IOU required
- IOU may be more complex than CPE
- o POLLING ADDS INTERFERENCE
 - Memory access
 - Program execution
- NOT COMPATIBLE WITH REAL TIME PROCESSES
 - Rate selection critical
- o BUT HARDWARE INTERFACE IS SIMPLE

✓ INTERMODULE COMMUNICATIONS

SUGGESTED IMPROVEMENTS

o FUNCTIONAL CHARACTERISTICS

- CPU to CPU Single - if attention

Stack - if polling

- CPU to IOP Stack Global
- IOP to CPU Stack Global
- IOP to IOP Single if attention

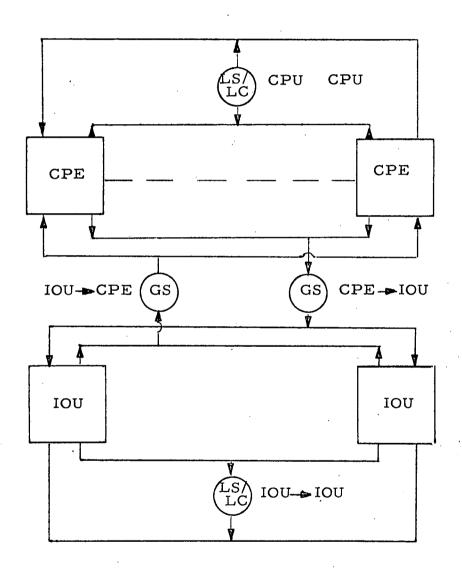
Stack - if polling

o IMPLEMENTATION CHARACTERISTICS

- All info. through memory
- External attention lines

INTERMODULE COMMUNICATION

SUGGESTED IMPROVEMENTS



/ INFORMATION PROTECTION

RCA DESIGN

- CONTENTS ERROR PROTECT
 - 11 bit ECC in storage
 - Burst ECC for Bus traffic
- ADDRESS ERROR PROTECT
 - Task page translate list
 - Protect key
- o SHARED ACCESS PROTECT
 - Lock and Go for shared instructions
 - Word status bits for shared data

Note: Currently planned to be software function

INFORMATION PROTECTION

RCA DESIGN

(continued)

o GENERAL COMMENTS

- Usage details on AEP and SAP not fully defined.
- Software interrogation of word status bits undesirable.
- AEP loses power for large pages/storage blocks.

SUGGESTED IMPROVEMENTS

- Expand LOCK AND GO to include shared data.

 Lock word can indicate instructions or data.
- Consider base/bound scheme for protection.
 (Writeup attached)
 - 1) Independent of page size
 - 2) Variable size blocks

Note: There may be a distinction between Control Exec. protect mechanisms and application program protect mechanisms.

FAULT TOLERANCE

RCA DESIGN

o FEATURES

- Single error correction
- Backup copies for critical info. (implementation undefined)
- Module faults limited to module in error (only if error detected in time)

GENERAL COMMENTS

- Objectives should be defined in detail unless not important
- Single instruction retry should be part of objectives

MIT - Hopkins, Fault Tolerant Processor

IBM - System 370

HUGHES - ARMMS

INTERMETRICS - Space Station MP

✓ POTENTIAL SUMC DESIGN IMPACTS

- o Paging related look-ahead
- o Interface with ATU
- o Interface with memory
- o Interface with channels/interrupts
- o Attention interface
- o Byte transform unit
- o Lock and go
- o Base/bound protection

REFERENCES

- 1) "Modular Space Station Computer Study", IBM Corporation, 27 October 1971, IBM No. 71W-00345.
- 2) "Space Ultrareliable Modular Computer SUMC Status Study", IBM Corporation, 28 August 1972. Bound copy of viewgraphs.
- 3) "SUMC Multi-Processor System Study", RCA Corporation, May 1972, Contract NAS12-2233, Contract Modification 7.
- ''Spaceborne Computer Executive Railine Functional Design Specification Volume II: Computer Executive Design for Space Station/Base''. Kennedy and Fitzpatrick, Contract NAS8-24930, NASA Report CR-1868. Computer Sciences Corporation, October 1971.
- 5) "System Configuration and Executive Requirements Specifications for Re-Usable Shuttle and Space Station/Base".

 Kennedy et al, Contract NAS8-24930, NASA Report CR-1820, Computer Sciences Corporation, May 1971.
- 6) "Multi-Processor Computer System Study". James S. Miller et al, Contract NAS9-9763, Intermetrics, Inc., March 1970.